DOCUMENT RESUME

ED 374 135 TM 021 990

AUTHOR Paulu, Nancy
TITLE Improving Math and Science Assessment. Report on the Secretary's Third Conference on Mathematics and Science Education (Washington, DC, September 20-21, 1993).
INSTITUTION Office of Educational Research and Improvement (ED), Washington, DC. Programs for the Improvement of Practice.
REPORT NO ISBN-0-16-045061-6; PIP-94-1201
PUB DATE Jun 94
NOTE 50p.
AVAILABLE FROM U.S. Government Printing Office, Superintendent of Documents, Mail Stop: SSOP, Washington, DC 20402-9328.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Collected Works - Conference Proceedings (021)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Curriculum Development; *Educational Assessment; Educational Change; Educational Improvement; Educational Policy; Educational Research; Elementary Secondary Education; *Mathematics; Parent Role; Policy Formation; Principals; *Sciences; Standards; *Student Evaluation; Student Role; Teacher Role; *Test Use
IDENTIFIERS Alternative Assessment; *Performance Based Evaluation; Reform Efforts

ABSTRACT

The third conference on improving mathematics and science assessment brought together more than 550 educators, researchers, and policymakers. Recommendations and information from the conference are incorporated into this report. Conference consensus was that students must recognize that there is more to assessment than grades and scores and begin to see assessment as a crucial and integral part of education. They should assume increasing responsibility for documenting and reflecting on their own progress as they move through school. Parents and the public must hold high expectations for instruction and for student participation. Teachers will hold all students, regardless of their backgrounds, to high standards and will recognize that assessment cannot be separated from learning. Principals and other administrators must oversee the reform of mathematics and science assessment and ensure that the curriculum reflects what we want students to know. An 18-item reading list is attached, and an appendix lists conference participants and sessions. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Improving Math and Science Assessment

Report on the Secretary's Third Conference on Mathematics and Science Education

Nancy Paulu
Programs for the Improvement of Practice

Office of Educational
Research and Improvement
U.S. Department of Education

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
Sharon P. Robinson
Assistant Secretary

Programs for the Improvement of Practice
Eve M. Bither
Director

June 1994

For sale by the U.S. Government Printing Office
Superintendent of Documents, Mail Stop: SSOP, Washington, DC 20402-9328
ISBN 0-16-045061-6

Contents

Foreword v

Recommendations and Responsibilities 1

Why Assessment Needs to Change 7

A New Assessment System—What We Need 11

Guidelines for Developing New Assessment Systems 17

Major Issues in Reforming Assessment 23

Where Do We Go From Here? 27

Endnotes 29

Reading List 31

Appendix 33

Acknowledgments 47

"To meet the challenges posed by a global economy, we must look to where the future is moving, not remain stuck where we are. Continuing achievement in the classroom means demanding more from students, setting higher standards, looking at what job skills they will need in tomorrow's economy, and offering them the opportunity to succeed. It requires keeping a focus on the goals of educational quality and community betterment."

Richard W. Riley
Secretary
U.S. Department of Education

Foreword

Throughout America, national efforts to improve achievement in mathematics and science education have moved slowly. Despite modest gains, national assessments show that most American children still do not understand and perform at levels needed for success in today's world. Many mathematics and science classes fail to challenge students to work to their full potential. Without higher performance standards and significant academic improvement, our nation cannot expect to meet the national education goal that cails for American students to be first in the world in mathematics and science achievement by the year 2000.

Toward this end, the U.S. Department of Education has sponsored three conferences in the 1990s that together aimed to improve mathematics and science education. The first focused on reforming the mathematics and science curriculum and on the need for higher standards. The second addressed the need to improve the mathematics and science preparation of America's teachers and on providing updated hands-on teaching materials.

The third conference on improving mathematics and science assessment is the subject of this report. On September 20-21, 1993, this conference drew together more than 550 educators, researchers, and policymakers to share insights about assessment and to learn from one another. Among the questions addressed were: Why must assessment change? What forms of math and science assessments can help American students succeed in these subjects? How can districts reforming assessment assure that tests are fair for students of all races and income levels and of both genders? What might it cost to reform assessment? How can better assessments fuel the drive toward comprehensive reform of American education and higher academic standards?

Recommendations and information from the 2-day conference are incorporated into this report. In forming answers to key questions about assessment, it is important to remember that no one test can tell us all we need to know about student achievement. A variety of tests and assessment systems, each making a different contribution, is needed to gauge student progress effectively.

For example, national assessments in mathematics and science can track progress toward the national math and science goal and be used to gather information comparing American students with those in other countries. State assessments can be used to gauge student progress toward statewide curricular goals. Schoolwide tests can measure progress toward district goals.

Within individual classrooms, tests can be used for what are perhaps their most important purposes—to help teachers discover what students know and are able to do and, subsequently, to use this information to guide instruction. Students also benefit by knowing what areas need more work. Assessments need to be viewed not as an unpleasant interruption to Friday afternoon, but as an integral part of the curriculum and of instruction.

Assessment reform alone is not a solution to all of America's educational challenges. Combined with a stronger curriculum and better teaching, however, it can move us toward comprehensive change and higher standards for all students.

Much remains to be learned about making assessments meaningful. By gathering together so many prominent leaders, the conference was an important step toward developing better assessment systems. We hope this report will alert still more educators, policymakers, and parents to the need for better ways to measure student achievement in mathematics and science so that together we can achieve this goal more rapidly.

Sharon P. Robinson
Assistant Secretary
Office of Educational Research and Improvement

Recommendations and Responsibilities

Participants at the Secretary's conference on improving assessment in mathematics and science education possessed a range of professional backgrounds and perspectives. During their discussions and presentations, many critical issues confronting assessment reformers emerged. Finding solutions to these issues will require the cooperation of many groups and individuals.

Listed here are recommendations and responsibilities for all whose contributions are needed to reform the assessment of math and science: students, parents, the public, teachers, principals and other education administrators, policymakers, teacher educators and schools of education, and test publishers and developers.

It is important that these recommendations be set in motion simultaneously because they are interconnected. Implementing one recommendation will have repercussions for others, and failing to implement one of them may impede efforts in another sphere of education reform.

After the recommendations and responsibilities, the remainder of this report provides an overview of the conference from which these recommendations and responsibilities are drawn.

Students

- Students will recognize that there is more to assessment than grades and test scores. Students will view assessment as a critical and integral part of education that can help them and their teachers identify areas to which they must devote more effort.

- As they move through school, students will assume growing responsibility for documenting and reflecting on their own progress. By high school graduation, students will possess the basic skills needed to assess their own work. Such skills will be required so they can continue to learn and evaluate themselves throughout their lives.

- Students will hold themselves to high standards and recognize that, whatever their background, they can do well on tests, in school, and in life.

Parents

- Parents, rich and poor alike, will hold high expectations for their children's mathematics and science learning and make clear to their children that they are expected to learn as much as their effort and ability will permit.

- Parents will understand that no one test score can provide an adequate indication of a youngster's capabilities. To gain a complete and current picture of what their child knows and can do, parents will look to: scores on a range of tests and assessments, grades, and the actual content of their child's school work. Parents will also gain critical information by communicating regularly with their child's teacher(s).

The Public

- All Americans, including those who have no school-aged children, will understand the importance of holding all students to high academic standards. The public will recognize that academic proficiency can contribute to success in the workplace and therefore can contribute to a strong national economy. Proficiency can also contribute to personal fulfillment and a better quality life.

- The public will become more aware of the advantages of using a variety of assessment techniques and tests to measure progress toward high performance standards.

Teachers

- Teachers will hold all students to high standards, regardless of their cultural, ethnic, or economic background or their gender.

- Teachers will recognize that assessment cannot be separated from learning and will weave assessment into their teaching.

- Teachers, working with their students, will assume primary responsibility for monitoring progress toward high standards. Teachers will provide students with extensive feedback on tests so that they can learn from their mistakes. Teachers will also use what they learn from assessment results to alter instruction for individual students as well as for groups of students or entire classes.

- Teachers will align assessment with the curriculum so that students are being tested on the material they are taught. This alignment will require teachers to be actively involved in evaluating and improving the curriculum so that it continues to reflect the challenging standards that students must meet to succeed in today's world.

- Teachers must help plan and participate in professional development activities that provide them with up-to-date information on assessment and enable them to merge assessment and learning successfully.

Principals and Other Education Administrators

- Principals and other education administrators will oversee the reform of mathematics and science assessment. They will make certain that the curriculum reflects what we want our students to know and be able to do, and they will take steps to assure that the assessment system is consistent with the curriculum.

- Principals and other education administrators will see that resources are available so teachers and other district educators can receive current information and training on using new forms of assessment.

- Principals and other education administrators will ensure that an appropriate balance is struck between tests (which are used largely for accountability) and assessments (which are used primarily to improve instruction).

- Principals and other education administrators must recognize the strengths and limitations of varying forms of assessment and take steps to assure that each assessment is used for its intended purpose.

- Principals and other education administrators will make certain that new forms of assessment, as well as existing ones, are not used to discriminate against any cultural or ethnic group or individual, or against students of either gender.

Policymakers

- Policymakers will provide leadership for evaluating assessment systems and take steps to improve the systems as needed. The evaluation must take place at all organizational levels: federal, state, district, school, and classroom.

- Policymakers will increase public awareness of the demands of the new world economy and the kinds of knowledge and skills students will need to succeed in it.

- Policymakers will provide the resources needed to develop and implement improved forms of assessment.

- Policymakers will encourage and support ongoing efforts to provide meaningful international comparisons of student achievement. Policymakers will use the findings to help identify areas in which American students must improve to be competitive in international marketplaces and will encourage educators to make improvements in these areas.

Teacher Educators and Schools of Education

- Teacher educators must provide more and better instruction to teachers in using new forms of assessment. This is critical if teachers are to assume

ED 374 135

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

SCOPE OF INTEREST NOTICE

The ERIC Facility has assigned this document for processing to: TM

In our judgment, this document is also of interest to the Clearinghouses noted to the right. Indexing should reflect their special points of view. SE

BEST COPY AVAILABLE

TM 021990

more assessment responsibilities and link tests more closely with learning. The instruction will be provided at colleges and university-run teacher training programs before teachers begin their careers as well as periodically after they have begun their careers.

Test Publishers and Developers

- Test publishers and others involved in the development of tests will recognize that a range of tests and assessment systems are needed in order to provide a broad picture of student achievement. In addition to traditional multiple-choice tests, students must be allowed to demonstrate their knowledge through such formats as portfolios, hands-on experiments, and tests that provide open-ended questions.

- Test publishers and others involved in creating sound assessment systems will work aggressively to develop, evaluate, and continually improve assessments that are aligned with high standards.

Why Assessment Needs to Change

Assessment has the potential to serve as a powerful force in improving American education. But more often than not, the traditional testing strategies used in American classrooms fail to measure progress toward the goals we have set for the reform of mathematics and science education.

Ideally, we would like our tests to:

- Gauge the ability of our students to solve complex problems;
- Tell us whether our young people can generate hypotheses and test them out; and
- Tell us whether students can think critically and apply math and science to their everyday lives.

But most assessment systems are not integrated into instruction in a way that makes them valuable to education. We need assessments that are used to:

- Measure what students know, instead of what they do not know;
- Educate children, rather than sort them; and
- Improve schools, rather than rank them.

Realtors use test scores to identify the "best" schools as a selling point for houses. Newspapers rank schools and districts by their test scores. Superintendents can be fired for low test scores. School boards have even been known to issue mandates requiring all students to be above the national norm—or above the 50th percentile. [1]

The need to transform education in this country is widely recognized and considered to be a paramount issue of the day. But an awareness of the need to reevaluate and change the nature of assessments and the way they are

conducted is not so widespread. Sound assessment systems can have a significant and positive impact on teachers, students, and parents. But despite the potential benefits of assessments, most states, districts, and teachers still rely on tests that alone are insufficient indicators of student achievement: national standardized tests, tests developed by textbook manufacturers that accompany their products, and state-developed examinations.

By relying on these traditional assessment procedures, our students suffer. Current mainstream tests can provide some important information to guide educational policy and student learning. But these tests have limitations. Most often, they:

- Measure a narrow range of academic skills, but not the full range of ways in which we now know that people can be intelligent [2] and

- Test a limited range of student responses, primarily through multiple-choice questions, but not a student's conceptual depth of understanding or ability to solve problems.

Moreover, when the test scores have serious consequences, teachers feel pressure to teach to the tests. Researchers have found ample evidence that testing shapes instruction—and that tested areas are often taught at the expense of untested ones. This is an unfortunate practice if the tests themselves fail to measure what we want students to know and be able to do.

Twenty years ago, standardized tests were a more acceptable gauge of student achievement than now. But much has changed since most tests in use today were developed. Our mathematical and scientific knowledge has expanded, and what is taught in the classroom has increased accordingly. Our understanding of how students learn has also improved, requiring us to assess more rigorously for conceptual understanding and ability to analyze and solve problems. Unfortunately, few standardized tests provide meaningful measurements in these areas.

As Hugh Burkhardt, professor and former director of the Shell Centre in Mathematical Education at England's University of Nottingham, explained at the conference:

> "[People] are losing their jobs all over the world now because you can buy automata that are a lot cheaper and have no health care requirements. Society is looking for students who are flexible, adaptable, and self-propelled in their learning. U.S. economic prosperity depends on its population being thinkers."

Creative assessment strategies to identify and foster such thinking are surfacing in some states and school districts, but they are not yet widespread. Major changes in assessing students' performance in math and science will be needed in more American classrooms to develop students who can think critically.

A New Assessment System—What We Need

Assessment needs to change—beginning with our understanding of what it can do for our nation and its youngsters. It is no longer sufficient to regard assessment solely as a way to assign grades to students and determine whether educators are doing a good job. A new and broader view of assessment needs to recognize the many purposes that it can serve. [3]

- Some assessments provide information about large groups of students. These can be useful for those wishing to change policies that affect math or science education. For example, these assessments can:

  — Monitor the overall achievement of students in a school, a district, a state, or the nation;

  — Evaluate the effectiveness of a particular program or an approach to learning;

  — Provide the basis for making improvements in education at the school, district, state, or national levels; and

  — Provide guidance on how resources can be allocated more effectively.

- Other kinds of assessments can serve what is perhaps a system's most important mission—improving instruction within American classrooms. The assessment activities themselves can provide opportunities to help students grasp concepts and think analytically. Teachers can use assessments to:

  — Guide instruction and make it more effective by establishing what students know and how they learn it as instruction proceeds;

  — Impress upon students, school staff, and parents the expectations for learning; and

— Document each student's progress throughout the year, or as a student moves to the next level of education or into the workplace.

This broad range of assessment purposes must be accompanied by a wider range of measurement systems and techniques. It is next to impossible to score objectively everything a student does. Therefore, it is wise to consider a testing system that measures the range of student abilities and uses different tools to measure different skills. Below are some examples that districts might consider to provide a comprehensive assessment of student achievement:

- **Portfolios** consist of materials that students and teachers select to represent a student's best work. Portfolios can record many things, for instance, progress in a particular subject over a short or a long period of time, progress on long-term projects, or performance in group work.

- **Computers** can help track the process of learning and thinking and can interact with students. For example, computers can record how students learn with feedback, and students' abilities to deal with realistic situations and problems.

  Computers can be beneficial for both large-scale assessments and assessments within individual classrooms. The Educational Testing Service, for example, recently announced that students will use computers instead of paper-and-pencil tests to take the Graduate Record Examinations. This is expected to measure their knowledge more accurately and provide results more rapidly. Within a classroom, a computer-assisted assessment might help diagnose individual student learning difficulties as well as accomplishments and help the teacher make instruction more effective.

- **Videotapes** can record ongoing activities and explanations in rich detail. This makes it possible to determine such things as how students explain ideas and answer challenging questions and how students carry out tasks and perform experiments. With videos, for example, students can be shown complex experiments, then asked to evaluate the methods used and to interpret the results.

- **Hands-on tasks** can be used to ask students to show what they know by working with or manipulating concrete materials. Science students might be instructed to design and conduct an experiment to test their hypotheses by using scientific equipment or procedures. For example, they might use a microscope to analyze and interpret information on slides.

- **Group activities** can be helpful to assess some skills. In a science assessment, for example, group tasks are needed to see how well team members work cooperatively, communicate, and integrate team members' contributions into a final product.

- **Paper-and-pencil assessments** can continue to be used, as appropriate, to measure such things as students' knowledge of facts, concepts, and procedures; their ability to solve problems; and their ability to comprehend text.

- **Multiple-choice tests** can be expected to continue playing a role because their format is efficient and scores on some multiple-choice standardized tests correlate moderately well with subsequent academic performance as measured by grades. Improvements in some of these multiple-choice tests must be made, however, to bring them more in line with new curricular goals.

Many examples of promising alternatives to traditional assessment were in evidence at the conference. All of them are listed in the appendices, and a few are described below.

New York's Elementary Science Program Evaluation Test is helping to determine how much science elementary school students know and how to help them understand scientific concepts better. This statewide assessment system is closely tied to the New York State science syllabus.

The assessment includes a "mystery box" test item. Every student assessed is asked to examine a small rectangular box, which is wrapped in plain brown paper and taped shut, and to guess what is inside. "Do not open the box!" the instructions read. One of the cardinal rules, in fact, is never to open the box,

because an unopened box captures the spirit of what science is about; when scientists go to the middle of the Earth by seismology or to the Moon or the planets, they are looking at a "mystery box."

"Pick up the box," the instructions continue. "Something is inside. . . Tilt the box or shake it gently. Listen carefully as you do. Roll the box in another direction. Listen again." Students are then asked to answer some questions. For example, what is one kind of motion made by an object in your box? What is the shape of an object in your box? There are many acceptable answers to all of the questions.

Students are given 7 minutes to complete the mystery box task, but Douglas Reynolds, who oversaw the implementation of New York's new elementary science assessment program, told conference participants that students could be given 2 hours "and they would still be sitting there observing, tilting, and shaking."

In addition to the hands-on assessment items, the New York assessment program includes more traditional multiple-choice questions, which Reynolds notes also have advantages. Results from both traditional and nontraditional test items are used to flag areas in which students are weak—and to adjust the academic program accordingly.

The **New Standards Project** (NSP) based at the University of Pittsburgh is doing pioneering work in state-based assessments. Lauren Resnick, the project's Co-Director, told conference participants about a task that middle-school students performed last spring on an NSP pilot examination. This task is typical of many of the New Standards Project assessment tasks because it provides challenge, it is set in a reasonably realistic context, and it asks students to justify what they are doing:

Students were asked to design customized laces for sports shoes. Instead of having just short, medium, and long laces, the idea is to provide customers with exactly the length they need depending on the number of islets in their shoes and the size of their feet. To help with this task, test-takers were given a chart with information said to be from actual customers and were asked to write a response to their boss on how to customize the shoe laces. Students were asked to include in this response: (1) a table that would show for each number of islets in the shoe how long the laces should be, and (2) a rule that the customer could use to figure out the length based on the number of islets and the need for a bow. Students were also asked to express this rule in a formula and to explain why the formula made sense.

Vermont has been a leader in developing a portfolio assessment system in mathematics and in writing as an alternative to traditional state tests. The portfolio program in math consists of 5 to 7 pieces of work selected by the students. The portfolios help to reveal much about their creators' knowledge and accomplishments that more traditional forms of assessment cannot do. Two studies of Vermont's emerging portfolio system by the RAND Corporation show that the new assessments are beginning to have a positive effect on curriculum and instruction but that the scoring of the portfolios is not yet reliable enough to be used for accountability purposes. However, a survey of teachers and interviews with principals revealed enthusiasm for the new system. A majority of the math teachers surveyed reported they were "frequently" more enthusiastic about teaching math, and 80 percent indicated that the portfolios had led them to change their estimation of student ability.

Robert Kenney, mathematics consultant for the Vermont Department of Education, stressed the importance of involving teachers in the assessments. "The more they are involved," he said, "the more they value new assessment as an instructional device."

Kentucky's new assessment program is a critical part of the state's Education Reform Act of 1990, which was passed in an effort to raise academic standards and achievement for all Kentucky students. For years, Kentucky schools administered a national standardized achievement test that benefited students who excel in remembering facts but was unable to measure other abilities (such as creativity, analyzing data, or working with a team to solve real-life problems).

Kentucky's new assessment program recognizes that children learn and apply learning in different ways. The new assessments, administered to children in grades 4, 8, and 12, include: multiple-choice and short-essay questions; performance tasks that call for students to solve simulated, real-life problems; and portfolios that present each student's best work collected throughout the year. Assessments are, or eventually will be, conducted in seven areas, including math and science.

Each student's scores place him or her into one of four performance categories: novice, apprentice, proficient, or distinguished. The assessment measures not only every student's progress but every school's success. Schools are evaluated on the progress students make from one testing period to another, and schools that do well are rewarded. Kentucky is phasing in the program, which began in the 1991-92 school year, over 5 years and is making adjustments as needed.

Guidelines for Developing New Assessment Systems

Mathematics and science education are in a period of transition. The curriculum, teaching, and assessment in these fields are all undergoing fundamental change. Because our knowledge of assessment continues to evolve, it is often hard to know whether reform efforts are on the right track. The following checklist can serve as a conceptual guide for designing new assessment systems.

- **Assessments must be coupled with higher performance standards.** To accomplish what we want, assessment programs must be designed to encourage learning—not just to measure it. This requires that higher academic standards be set at all institutional levels—local, state, and federal—and that assessment systems reflect these higher standards. It requires, too, that students devote sufficient time to studying what we want them to know and be able to do. As Emerson J. Elliott, Commissioner of the National Center for Education Statistics, explained at the conference, international studies comparing student performance show that "the more students are taught, the more they learn, and the better they perform on tests."

  Toward this end, the National Council of Teachers of Mathematics (NCTM) has developed math curriculum standards which are in use around the country. And the National Academy of Sciences, with a grant from the U.S. Department of Education's Office of Educational Research and Improvement and with support from a coalition of other federal agencies, is developing science standards it hopes will guide what science is taught, as well as how it is taught and assessed. Local and state officials in the process of setting new achievement standards can look to these voluntary national standards for guidance in the future.

  To achieve higher standards, everyone involved in reforming education must assume that *all* students, whatever their cultural and

economic background, can learn complex math and science. Curriculum and teaching practices that encourage students to think and solve complex problems must be adopted, and tests must be developed that can help gauge student response to these new practices.

- **Assessment systems must measure what we value as opposed to what is easy to measure.** In the area of mathematics, NCTM is urging that reform—and assessment—be built around a drive to help students develop "mathematical power." Thomas A. Romberg, Director of the National Center for Research in Mathematical Sciences Education at the University of Wisconsin in Madison, explained this notion at the conference:

  > "All students need to achieve and to become problem solvers, to value mathematics, to reason and communicate mathematically, and to be competent at using mathematics to make sense of real-world problem situations."

  This notion of mathematical power requires that students be taught both math facts and topics *and* how to apply them in daily activities. What we value must also be reflected in assessment plans, in each assessment task and activity, in the way students' responses are interpreted and in the way results are used. [4]

  In the area of science, we want students to understand the major concepts in the sciences and that the science disciplines involve more than facts and principles—that they also involve methods and processes of inquiry. We want students to be able to think like scientists. We want students who can identify useful information, interpret it, make inferences using scientific information, make predictions, and apply their scientific knowledge to new situations. [5]

  Current science assessment seldom measures all of these abilities. To move toward better assessments in science, Audrey Champagne, professor of chemistry and Chair of the Department of Education Theory and Practice at the State University of New York in Albany,

suggested at the conference that we begin by determining how much science students understand. An explanation—why the student gave a particular answer—is often more important than the specific answer. She provided the following example:

> To assess a student's understanding about evaporation, we might imagine a situation where she is asked to figure out which of several methods will dry a soccer jersey most rapidly. Two possibilities are to ask the child to select the fastest method, or to order the methods from fastest to slowest. But we'll develop a better understanding of what the child knows by asking why she has placed the methods for drying in a particular sequence.

We must also assess a student's ability to demonstrate an understanding of the philosophy of science and use a scientific perspective to evaluate the quality of information. For example, a high school graduate should know that an article about nutrition in the *New York Times* which cites facts from a report from the National Academy of Sciences would be more reliable than a short item about nutrition without supporting information in the *National Enquirer*.

- **Assessment should help students learn mathematics and science.** This requires that assessments: be woven into instruction; be used to engage students in mathematical and scientific tasks; be based on what is known about how students learn these subjects; and provide opportunities for students to evaluate, reflect upon, and improve their work.

  The National Center for Improving Science Education has devised questions that teachers and administrators should ask when they evaluate the quality of an elementary school science test. [6] For example:

  - Do some of the problems require students to think about and analyze situations?

- Does the test feature sets of problems that call for more than one step in arriving at a solution?

- Are students encouraged to use a variety of approaches to solve a problem?

- Do students have opportunities to make up their own questions, problems, or designs?

■ **Assessments must be equitable.** They must be reliable and valid indicators of what *all* students know and have an opportunity to learn. They should help *all* students—in inner-city schools, in suburban schools, and in rural areas—demonstrate their knowledge of math and science and encourage each student to meet the high expectations set in these subject areas. Boys and girls must benefit equally from assessments, as must students from different racial and ethnic groups. This requires that the assessment process provide all students with appropriate and varied ways to show what they know, and take into account students' perspectives, reasoning patterns, and assumptions.

■ **Every aspect of an assessment system, including its design, should be consistent with its purpose.** Assessment activities must reflect what we want to assess; we must not use a computation test when we need one that measures a student's ability to solve problems. While this may seem obvious, the NCTM reports that such a lack of consistency is perhaps the most widely violated principle in educational testing practice.

■ **Teachers must be actively involved in reforming assessment and in assessing students.** Champagne explains:

> "In terms of *who* should assess, clearly assessment should be in the hands of those who can do the best job of assessing, as well as those who can do the best job of making use of the information from the assessment. That turns out to be the teachers and the students."

In most assessment systems today, teachers play a minor role in developing or refining tests, and in even fewer schools are tests used constructively to guide instruction. As assessment is reformed, this role must be expanded.

The teacher's ultimate goal must be to help students learn to assess the quality of their own work and to become responsible for their own progress. These abilities will be critical to students in many activities and settings both at home and in the workplace.

- **New assessments must be open to review and scrutiny.** Everyone affected by new assessment needs information about it. Teachers need to understand its goals and how it will be administered. Students must be told what they need to know, how they will be expected to demonstrate it, and what the consequences of the assessment will be.

Major Issues in Reforming Assessment

Even with the most promising models of assessment systems and the active involvement of many educators, it is a challenge to construct forms of assessment that test what we want children to know. "It's not easy to be innovative, consistent, and fair," Madeleine Kunin, Deputy Secretary for the U.S. Department of Education, told conference participants. Reforming assessment in mathematics and science requires that the following issues be addressed:

- **The American mind-set.** Some educators, policymakers, and parents hold some outmoded assumptions concerning assessment. First, many continue to view tests primarily as a tool to "grade" education systems. We must work hard to communicate that *assessments can and must play a central role in helping students to learn.*

  Second, many hold the view that one number—the score that a student receives on one measurement like the IQ test or the Scholastic Aptitude Test, for example—tells the whole story. For reform to succeed, we need to communicate that different tests tell us different things, and that a variety of measurements together can provide a more complete picture of our children's abilities. Champagne explains:

  > "People in the education community have raised a very large population of individuals who believe in standardized tests and who believe that numbers are the best descriptors of students' understanding . . . A major job is to change that public perception, and to make the public a little less confident in standardized tests and a bit more confident in what their children's science and mathematics teachers can tell them about the student's understanding and progress."

  Changing the mind-set also involves developing public confidence in the ability of teachers to use assessments to make sound educational

decisions. Many parents and policymakers are uncomfortable making teachers the primary assessors and shy away from giving them additional assessment responsibilities. Any meaningful change in assessment, however, is not likely to take place unless teachers are more actively involved in making them happen. Assessment helps teachers monitor the success of their teaching and make better instructional decisions.

- **Teacher preparation.** Teachers must have more and better preparation in using new forms of assessment. This is the key to changing the American mind-set. Unfortunately, most teacher education programs fail to provide adequate training in any form of assessment, and most classroom teachers adhere to very traditional assessment practices. Districts and schools must work together with teachers to help them link assessments with instruction and use the test results effectively.

- **Equity.** Issues related to equity and fairness rest at the heart of concerns being raised about assessment reform, according to Marshall Smith, Under Secretary in the U.S. Department of Education. Parents and policymakers want to be confident that children from all economic and cultural backgrounds, as well as students of both sexes, have an opportunity to learn the knowledge and skills that will be tested, and that all students have access to assessment systems that are integrated with opportunities to learn. Finally, because alternative forms of assessment that call for judgments to be made afford greater risks for subjective evaluations, steps must be taken to assure that results are not skewed against any group or individual student.

- **Cost.** The cost of new assessments is higher in time and in money than the cost of more traditional ones. Nontraditional assessment materials are more expensive, and more time is required to train teachers and other test administrators to use the materials. Moreover, performance-based tests require more time to develop, take, score, and interpret—all of which adds to the price tag. (Performance-based tests require students to apply their knowledge in real-world situations.) A recent General Accounting Office report on the subject found that multiple-choice tests averaged less

than half the cost per student of performance-based tests—$16 versus $33 per student annually. [7]

To hold down costs, we need models of nontraditional assessment systems so that individual states and districts do not duplicate each other's efforts. New assessment systems must also be phased in to spread out the expenses.

Where Do We Go From Here?

The issues discussed in this report are being addressed further during regional conferences and activities on the reform of mathematics and science assessment. The U.S. Department of Education has asked each of its 10 regional education laboratories to organize post-conference activities for reforming assessment in order to sustain the momentum achieved at the national conference. The regional activities will enable issues raised at the national conference to be discussed in more detail and lay the foundation for everyone involved in reform efforts to engage in more in-depth work.

For major change to take place in the way we assess student achievement in mathematics and science, the commitment of the American people is required. Guidance is needed from Washington, DC, and responsive oversight is needed from state capitals. But commitment must also come from homes, from neighborhoods, and from local communities. Educators, researchers, policymakers, and parents together must initiate ideas for improving assessment and must work together to bring about the needed changes.

More probing assessments alone will not improve mathematics and science education. But when combined with a revitalized curriculum and more effective teaching, reform can be set in motion. Assistant Secretary Sharon Robinson explained:

> "...what we want to achieve is nothing less than a system reformed to focus on a child's strengths, one that brings the child forward with optimism about her or his potential to learn, and one that reveals at the same time both the necessity and pleasure of learning as a life-time pursuit."

These are difficult tasks, but our children deserve our best efforts.

Endnotes

1. See Shephard, Lorrie A., "Why We Need Better Assessments," in *Educational Leadership*, April 1989.

2. See Gardner, Howard, *Frames of Mind: The Theory of Multiple Intelligences*, Basic Books, Inc., New York, 1983.

3. See the National Center for Improving Science Education's *The High Stakes of High School Science*, pp. 34-57, 1991.

4. See "Highlights from Assessment Standards for School Mathematics," prepared by the Working Groups of the National Council of Teachers of Mathematics, October 1993.

5. See *Measuring What Counts*, a policy brief by the Mathematical Sciences Education Board and the National Research Council, National Academy Press, Washington, DC, 1993.

6. See *National Science Education Standards*, a working paper of the National Committee on Science Education Standards and Assessment, National Research Council, Washington, DC, July 1993.

7. See *Student Testing: Current Extent and Expenditures, With Cost Estimates for a National Examination*, prepared by the United States General Accounting Office, Washington, DC, January 1993.

Reading List

1. Baker, E.L., O'Neil, H.F. Jr., and Linn, R.L., "Policy and Validity Prospects for Performance-Based Assessment," (1989). *American Psychologist*, 48(12), Pages 1210-1218.

2. Barton, Paul. *National Standards for Education: What Might They Look Like; A Workbook*, (1992). Available from: Educational Testing Service, Policy Information Center, Mail Stop 04-R, Princeton, NJ 08541.

3. Brandt, Ron, "Using Performance Assessment." *Educational Leadership*, 49(8), 1992. Entire issue.

4. Champagne, Audrey B., Lovitts, B.E., and Calinger, B.J., *Assessment in the Service of Instruction*, (1990). Available from: American Association for the Advancement of Science, 1933 H Street NW, Washington, DC 20005.

5. Charles, R., and Silver, E., editors, *The Teaching and Assessing of Mathematical Problem Solving*, (1988). National Council of Teachers of Mathematics.

6. Commission on Standards for School Mathematics, *Curriculum and Evaluation Standards for School Mathematics*, (1989). Available from: National Council of Teachers of Mathematics, 1906 Association Drive, Reston, VA 22091.

7. EQUALS. *Assessment Alternatives in Mathematics*, (1989). Available from: University of California, Lawrence Hall of Science, Berkeley, CA 94720.

8. Hein, George, editor, *The Assessment of Hands-On Elementary Science Programs* (1990). Grand Forks, ND: University of North Dakota.

9. Herman, J.L., Aschbacher, P.R., and Winters, L., *A Practical Guide to Alternative Assessment*, (1992). Association for Supervision and Curriculum Development, Alexandria, VA.

10. Kulm, Gerald, *Assessing Higher Order Thinking in Mathematics*, (1990). Available from: American Association for the Advancement of Science, 1333 H Street NW, Washington, DC 20005.

11. National Center for Improving Science Education, *The High Stakes of High School Science*, (1991). The Network, Inc., Andover, MA.

12. Raizen, Senta, Baron, J.B., Champagne, A.B., Haertel, E., Mullis, I.V.S., and Oakes, J., *Assessment in Elementary School Science Education*, (1989). Report of the National Center for Improving Science Education, The Network, Inc., Andover, MA.

13. Romberg, Thomas, *Mathematics Assessment and Evaluation: Imperatives for Mathematics Educators*, (1992). Available from: State University of New York Press, State University Plaza, Albany, NY 12246.

14. Sternberg, Robert J., "Thinking Styles: Keys to Understanding Student Performance, *Phi Delta Kappan*, 71, Pages 366-371, January 1990.

15. Sweet, David, (1993). "Performance Assessment." Number 2 of *Education Consumer Guide*. Office of Educational Research and Improvement, U.S. Department of Education, Washington, DC.

16. Sweet, David, (1993). "Student Portfolios: Classroom Uses." Number 8 of *Education Consumer Guide*. Office of Educational Research and Improvement, U.S. Department of Education, Washington, DC.

17. Sweet, David, (1993). "Student Portfolios: Administrative Uses." Number 9 of *Education Consumer Guide*. Office of Educational Research and Improvement, U.S. Department of Education, Washington, DC.

18. Wiggens, G.P., *Assessing Student Performance*, (1993). Jossey-Bass Publishers, San Francisco, CA.

Appendix

Presenters and Demonstrators

The Department is indebted to the following individuals who made presentations, moderated panels, or led demonstration sessions:

Presenters and Moderators

Gordon Ambach
Executive Director
Council of Chief State School Officers
Suite 700
One Massachusetts Avenue NW
Washington, DC 20001-1431

Eva Baker
Co-Director
Center for Research on Evaluation, Standards, and Student Testing
University of California - Los Angeles
Maxxam Building
10880 Wilshire Boulevard
Los Angeles, CA 90024-1522

Eve M. Bither
Director
Programs for the Improvement of Practice
Office of Educational Research and Improvement
U.S. Department of Education
Room 500E
555 New Jersey Avenue NW
Washington, DC 20208-5572

Linda Bryant
Dean of Students
Florence Reizenstein Middle School
129 Denniston Avenue
Pittsburgh, PA 15206

Hugh Burkhardt
Director
Shell Centre for Mathematics Education
Nottingham University
Nottingham, United Kingdom
NG7 2RD

Michelle Cambier-Manon
Director of Bilingual Science
Community School District #6
406 East Second Street
Brooklyn, NY 11218

Audrey B. Champagne
Professor of Chemistry and Chair
Department of Education Theory and Practice
ED127A
State University of New York
Albany, NY 12222

Carlos R. Cisneros
State Senator and Chairman
Legislative Education Study Committee
Santa Fe, NM 87503

Michael Cohen
Senior Advisor to the Secretary
U.S. Department of Education
Room 4181C
400 Maryland Avenue SW
Washington, DC 20202-0100

Nancy Cole
President
Educational Testing Service
Rosedale and Carter Roads
Princeton, NJ 08541

Joseph Conaty
Acting Director
Office of Research
Office of Educational Research and Improvement
U.S. Department of Education
Room 610D
555 New Jersey Avenue NW
Washington, DC 20208-5573

Margaret Cozzens
Director
Division of Elementary, Secondary, and Informal Education
National Science Foundation
Room 885
4201 Wilson Boulevard
Arlington, VA 22230

Emerson J. Elliott
Commissioner
National Center for Education Statistics
Office of Educational Research and Improvement
U.S. Department of Education
Room 400F
555 New Jersey Avenue NW
Washington, DC 20208-5574

Richard Hill
President
Advanced Systems in
Measurement and
Evaluation, Inc.
P.O. Box 1217
Dover, NH 03820

Kenneth Hoffman
Associate Executive Officer for
Education
National Research Council
2101 Constitution Avenue NW
Harris Building, Room 450
Washington, DC 20418

Vinetta Jones
Director
Equity Programs
The College Board
45 Columbus Avenue
New York, NY10023-6992

Madeleine Kunin
Deputy Secretary
U.S. Department of Education
Room 4015
400 Maryland Avenue SW
Washington, DC 20202-0500

Douglas Lapp
Executive Director
National Science Resources
Center
Smithsonian Institution
Room 1201
Arts and Industries Building
Washington, DC 20560

Shirley Malcom
Head
Directorate for Education
and Human Resources
American Association for the
Advancement of Science
1333 H Street NW
Washington, DC 20005

Barbara R. Morgan
Teacher in Space Designee
c/o NASA Headquarters
ATTN: Pam Bacon
Code FET
300 E Street SW
Washington, DC 20546

Frederick Mulhauser
Assistant Director
Program Evaluation and
Methodology Division
Room 5735
U.S. General Accounting Office
441 G Street NW
Washington, DC 20548

E. Benjamin Nelson
Governor of Nebraska
Lincoln, NE 68509

Thomas Payzant
Assistant Secretary for
Elementary and Secondary
Education
U.S. Department of Education
Portals Building, Room 4000
1250 Maryland Avenue SW
Washington, DC 20202-6100

Lauren Resnick
Co-Director, The New Standards
Project, and Co-Director, The
Learning Research and
Development Center
824 LRDC Building
University of Pittsburgh
Pittsburgh, PA 15260

Douglas Reynolds
Consultant to the Northeast and
Islands Eisenhower Mathematics
and Science Consortium
Suite 900
300 Brickstone Square
Andover, MA 10810

Richard W. Riley
Secretary
U.S. Department of Education
Room 4181
400 Maryland Avenue SW
Washington, DC 20202-0100

Sharon P. Robinson
Assistant Secretary for
Educational Research and
Improvement
U.S. Department of Education
Room 602
555 New Jersey Avenue NW
Washington, DC 20208-5530

Thomas Romberg
Director
National Center for Research in
Mathematical Sciences Education
University of Wisconsin, Madison
Wisconsin Center for Education
Research
1025 West Johnson Street
Madison, WI 53706

Marshall Smith
Under Secretary
U.S. Department of Education
Room 4169
400 Maryland Avenue SW
Washington, DC 20202-4100

Robert J. Sternberg
IBM Professor of Psychology and
Education
Department of Psychology
Yale University
P.O. Box 208205
New Haven, CT 06520-8205

Dennie Palmer Wolf
Executive Director
PACE (Performance Assessment
Collaborative for Education)
Harvard University
8 Story Street
Cambridge, MA 02138

Demonstration Sessions

(Demonstration titles are in bold.)

1. An Urban District Looks at Assessment Reform (San Diego, CA)

Ruben A. Carriedo, Assistant Superintendent
Linda Carstens, Specialist
Planning, Assessment, and Accountability Division
San Diego City Schools
4100 Normal Street
San Diego, CA 92103

This session focused on San Diego's efforts to increase equity and opportunity for a diverse body of students by reforming assessment.

2. The Hardest Part of Performance Assessment: Experience from Connecticut's Common Core of Learning Assessment in High School Science and Mathematics

Joan Boykoff Baron
Educational Consultant
Division of Research, Evaluation and Assessment
State Department of Education
Box 2219
Hartford, CT 06145

This session described steps taken in Connecticut to reform assessment in high school mathematics and science. The session concentrated on the key role teachers played in developing, implementing, and scoring performance tasks.

3. Alternative Assessment in Math and Science: Virginia Teachers Focused on the Future

Jane Hange
Director for the Classroom Instruction Program
Appalachia Educational Laboratory
P.O. Box 1348
Charleston, WV 25325

Linda Hite
Teacher
3016 Chelsford Way
Williamsburg, VA 23185

Helen Rolfe
Director of Instruction and Professional Development
Virginia Education Association
116 South Third Street
Richmond, VA 23219

This session focused on Virginia's experience with teacher-development assessment tasks designed to measure the knowledge and skills students will need in "the real world."

4. "I put the washers in the cups and the shorter one is faster than the longer one."

David Anderson, Associate Director for Assessment
Katherine Pedersen, Project Director
South Dakota/National Science Foundation Statewide
Systemic Initiative
510 East Franklin Avenue
Pierre, SD 57501-3726

This session provided insights into South Dakota's efforts to design and implement a statewide performance assessment system.

5. Maryland's Systemic School Reform Initiative: The Implications of State Level Performance Assessments for Policymakers

Steven Ferrara, State Assessment Director
Robert E. Gabrys, Assistant State Superintendent for School Performance
Gary Heath, Science Coordinator
Maryland State Department of Education
200 West Baltimore Street
Baltimore, MD 21201

This session provided an overview of what it means for a state to include new performance assessment as part of its reform package. The session aimed to communicate that performance assessment alone is not a panacea for educational reform.

6. Measuring up to the Challenge of Assessment Reform

Linda P. Rosen
Associate Executive Director
Mathematical Sciences Education Board
Harris Building, Suite 476
2101 Constitution Avenue NW
Washington, DC 20418

This session discussed work currently under way in the mathematics education community to develop principles for assessment reform and to create examples of assessment tasks that bring these principles to life.

7. Performance Assessment as Part of a State Assessment Plan: The Experience of the Arizona Student Assessment Program

Paul Koehler
Associate Superintendent for Education Services
Arizona Department of Education
1535 West Jefferson
Phoenix, AZ 85007

This session provided information on the concept, value, construction, and use of performance-based assessments.

8. Pet Rocks

Susan Agruso
Coordinator
Curriculum and Assessment Services
State Education Department
Albany, NY 12234

This session provided information about the assessment system being developed in New York State. Participants in a sample assessment applied their conceptual understanding of Earth science to explain the characteristics and formation of a rock.

9. Education Reform in Kentucky: The Role of Alternative Assessment

Edward Reidy
Deputy Commissioner for Learning Results Services
1924 Capitol Plaza Tower
500 Metro Street
Frankfort, KY 40601

This session provided information about Kentucky's experience in implementing a statewide alternative assessment program.

10. Developing Sound Alternative Assessments: Tools from National R&D Efforts

Judith Arter, Senior Associate
Evaluation and Assessment
Northwest Regional Educational Laboratory
Suite 500
101 SW Main Street
Portland, OR 97204

Linda Ann Bond
Director of Assessment
Regional Policy Information Center
North Central Regional Educational Laboratory
Suite 300
1900 Spring Road
Oak Brook, IL 60521

Joan Herman
Associate Director
Center for Research on Evaluation,
Standards and Student Testing (CRESST)
University of California, Los Angeles
Maxxam Building
10880 Wilshire Boulevard
Los Angeles, CA 90024-1522

Ed Roeber
Director of Student Assessment Programs
Council of Chief State School Officers
Suite 700
One Massachusetts Avenue NW
Washington, DC 20001

Arie van der Ploeg
Database Supervisor
North Central Regional Educational Laboratory
Suite 300
1900 Spring Road
Oak Brook, IL 60521

This session provided information on local, state, and national efforts to develop alternative assessments in math and science. It also acquainted participants with resources for staff development that are available through the national research and development network of regional education laboratories and national education research centers.

11. Mathematics Portfolios in Vermont

Robert Kenney
Mathematics Consultant
Vermont Department of Education
120 State Street
Montpelier, VT 05602

This session focused on Vermont's efforts to use mathematics portfolios for instruction and assessment, a system developed by teachers as an alternative to traditional state tests.

12. California's Learning Assessment System: Performance-Based Accountability for Meaning-Centered Instruction

Sue Bennett
Administrator for Planning, Coordination and
Policy Services
California Department of Education
721 Capitol Mall, 4th Floor
P.O. Box 944272
Sacramento, CA 94244-2720

This session provided a look at California's efforts to encourage curriculum and instructional reform in math and science by developing an assessment system that reports achievement based on common statewide performance standards.

13. Mathematics Assessment: Beyond Multiple Choice

Timothy J. Crockett
Director of Program Management
Advanced Systems in Measurement and Evaluation, Inc.
P.O. Box 1217
Dover, NH 03820

This session focused on the types of information that can be learned from alternative assessment and the impact that assessment can have on instruction and the curriculum. The session introduced participants to the "open response" testing format.

14. Incorporating Alternative Mathematics Assessment into an On-going Assessment Program: One District's Experience

Diane J. Briars
Director, Division of Mathematics
Pittsburgh Public Schools
850 Boggs Avenue
Pittsburgh, PA 15211

This session focused on Pittsburgh's efforts to incorporate nontraditional assessment practices into its mathematics programs.

15. Pacesetter: Setting a Faster Pace for Reform

Elizabeth Badger
Director of Instructional Assessment Initiatives
The College Board
45 Columbus Avenue
New York, NY 10023

This session focused on the College Board's secondary school reform initiative—Pacesetter—a program that integrates instruction, teacher training and support, and assessments.

16. Algebridge and LearningPlus: Linking Assessment to Instruction

Paul Ramsey
Senior Development Leader
Educational Testing Service
Rosendale and Carter Roads
Princeton, NJ 08541

This session used two examples (Algebridge and LearningPlus) to show what assessment looks like when it is linked to instruction.

17. Assessment and Issues of Equity

Karen Mayfield, Jean Kerr Stenmark
Mathematics Specialists
EQUALS, FAMILY MATH and the Interactive Mathematics Programs
Lawrence Hall of Science
University of California
Berkeley, CA 94720

This session focused on the inequities fostered by some current modes of mathematics assessment and highlighted assessment systems that support curricula encouraging students to solve problems.

18. Integrated Learning: Creating Twenty-first Century Problem-solvers

Dennis McFaden, Barbara Nelson
Teachers
Thomas Jefferson High School for Science and Technology
6560 Braddock Road
Alexandria, VA 22312

This session focused on the need to teach the process of learning, rather than just information. It provided information about an interdisciplinary program in which students learn many things in addition to the core content material, including time management and how to use technology and to set goals. Such a program demands new forms of assessment.

19. NSTA Prototype Interactive Assessment Disc

Bill Aldridge, Executive Director
Carol Rudisill, Project Director of Assessment Materials
National Science Teachers Association
1742 Connecticut Avenue NW
Washington, DC 20009-1171

Julie Williams
Business Manager
Rocket Science Productions
3838 Folly Quarter Road
Ellicott City, MD 21042

This session focused on the use of a disc designed to test a student's performance and understanding of certain scientific concepts. The disk, developed by NSTA with Rocket Science Productions, requires minimal equipment.

Acknowledgments

A special thank you to those who contributed generously to this report. Joyce Stern, with able assistance from Kathy Fuller, organized the Secretary's conference on improving mathematics and science assessment. They both provided thoughtful suggestions and extensive background information for the report, and they reviewed drafts.

Barbara Lieb, Director of the Research Applications Division, and Eve Bither, Director of Programs for the Improvement of Practice, provided critical guidance. Others who gave advice and information include Eva Baker, Lance Ferderer, Joan Herman, Carol Lacampagne, Sharif Shakrani, Mary Sivertsen, Charles Stalford, and Joe Vaughan. Thanks also to Margery Martin for editing this report and to Phil Carr and Donna DiToto for designing the format and cover.

ISBN 0-16-045061-6

United States
Department of Education
Washington, DC 20208–5643

Official Business
Penalty for Private Use, $300

Postage and Fees Paid
U.S. Department of Education
Permit No. G–17

Third Class

PIP 94–1201